AD

U.S. ARMY WEAPONS COMMAND

AD 690185

TECHNICAL REPORT

WVT-6923

MAXIMUM LIKELIHOOD ESTIMATION OF DISTRIBUTION PARAMETERS FROM GENERALLY CENSORED SAMPLES

BY

ROYCE W. SOANES, JR

JUNE 1969

BENET R&E LABORATORIES

WATERVLIET ARSENAL

WATERVLIET—NEW YORK

DISPOSITION

Destroy this report when it is no longer needed. Do not return it to the originator.

DISCLAIMER

The findings in this report are not to be construed as an official Department of the Army position.

MAXIMUM LIKELIHOOD ESTIMATION OF DISTRIBUTION PARAMETERS FROM GENERALLY CENSORED SAMPLES

ABSTRACT

A simple derivation of the likelihood function for generally censored samples is given, where "generally censored sample" connotes: a set of failure data containing both failures and survivors with various different tentative survival times.

Cross-Reference Data

Maximum Likelihood

Likelihood Function

Censored Data

Truncated Tests

TABLE OF CONTENTS

1. INTRODUCTION

In an ideal situation, one is presented with a set of failure data in which every piece of data indicates that a failure has taken place, i.e., one has information on failures and on failures only. In this simplest of situations, parameter estimation may be carried out in a variety of relatively simple ways.

More often than not, however, this situation is not the case; instead, one is presented with a number of indicated failures and some additional information on items which have been tested to varying degrees, but not to failure. The reason for lack of further testing on the unfailed specimens may be an economic one or it may be that the unfailed items have been discarded for virtually any reason. Irrespective of the reasons involved, the data has been "censored", i.e., someone has either purposely or inadvertently prevented us from seeing just exactly when some of the items tested will fail. This report deals with the general method of maximum likelihood as applied to this general censoring problem.

2. MULTINOMIAL PROCESS

The multinomial process is of central importance in the derivation of the censored likelihood function. A multinomial process is simply the repeated performance of an experiment $\mathcal{E}$ n times where the experiment has r (mutually exclusive and exhaustive) outcomes with associated probabilities. A description of the process may be abbreviated symbolically:

Perform experiment $\mathcal{E}$ n times where the outcomes of $\mathcal{E}$ are:

$$\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_r$$

and

$$P(\mathcal{O}_i) = p_i \quad i = 1, 2, \ldots, r$$

Repetition of experiment $\mathcal{E}$ may be viewed as a composite experiment $\mathcal{E}_c$, a typical outcome of which is:

$$\mathcal{O}_c \equiv n_1\ \mathcal{O}_1\text{'s},\ n_2\ \mathcal{O}_2\text{'s},\ \ldots$$
$$n_r\ \mathcal{O}_r\text{'s}$$

are observed. The probability of this typical outcome of $\mathcal{E}_c$ is given by:

$$P(O_c) = P(n_1, n_2, \ldots, n_r)$$

$$= \frac{n!}{n_1! n_2! \ldots n_r!} p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r} \qquad (1)$$

This formula is given in any elementary statistics text under the heading "multinomial distribution." It should be noted that $P(O_c)$ is merely the probability of an <u>outcome</u> of E_c ; in general one wants the probability of an event (aggregation of outcomes). An event is given by a set of conditions on the n_i .

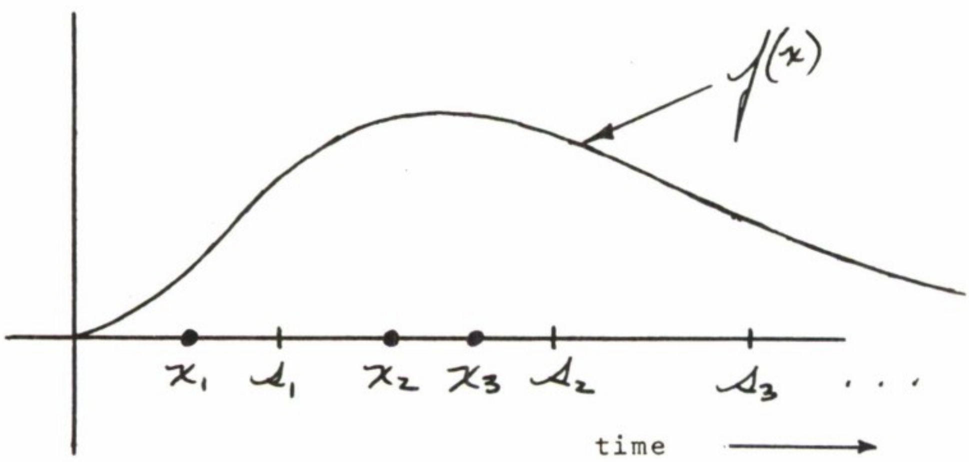
Therefore,

$$P(A) = \sum P(n_1, n_2, \ldots, n_r) \quad \text{(such that the } n_i \text{ satisfy condition A)} \qquad (2)$$

3. LIKELIHOOD FUNCTION

The principle of maximum likelihood dictates that we find the parameters which maximize the theoretical probability of obtaining the results we actually did obtain, i.e., other results "should" be less likely than the actual results.

The theoretical probability of obtaining the observed results must therefore be obtained first. The event describing the observed results is: K x's fail at specific times $x_1, x_2, \ldots, x_K$ and N x's survive for various lengths of time $s_1 \le s_2 \le \cdots \le s_N$ without failing $(N + K = n)$. Pictorially, the actual test results will look like:



A theoretical one by one test of each item to failure may be viewed as a multinomial process with the following outcomes (for $\mathcal{E}$).

$$\mathcal{O}_1' \equiv x \in (x_1, x_1 + dx)$$

$$\mathcal{O}_2' \equiv x \in (x_2, x_2 + dx)$$

$$\vdots$$

$$\mathcal{O}_K' \equiv x \in (x_K, x_K + dx)$$

$$\mathcal{O}_1 \equiv x \in (s_1, s_2)$$

$$\mathcal{O}_2 \equiv x \in (s_2, s_3)$$

$$\vdots$$

$$\mathcal{O}_{N-1} \equiv x \in (s_{N-1}, s_N)$$

$$\mathcal{O}_N \equiv x \in (s_N, \infty)$$

$\mathcal{O}_{N+1} \equiv x$ is elsewhere.

Strictly speaking, these outcomes are not mutually exclusive in a set theoretic sense; they are in a measure theoretic sense, however, and this is sufficient. This abuse of notation should be tolerated here for the sake of brevity.

If f is the underlying density function and F is the distribution function, the probabilities corresponding to these outcomes are:

$$p_1' = f(x_1)\,dx$$
$$p_2' = f(x_2)\,dx$$
$$\vdots$$
$$p_K' = f(x_K)\,dx$$
$$p_1 = F(s_2) - F(s_1)$$
$$p_2 = F(s_3) - F(s_2)$$
$$\vdots$$
$$p_{N-1} = F(s_N) - F(s_{N-1})$$
$$p_N = F(\infty) - F(s_N) = 1 - F(s_N)$$

Now that $\mathcal{E}$ has been specified, the probability of obtaining the observed results in a theoretical test of n items to failure is desired. The event describing the observed results is:

$$n_1' = 1$$
$$n_2' = 1$$
$$\vdots$$
$$n_K' = 1$$
$$n_1 = k_1$$
$$n_2 = k_2$$
$$\vdots$$
$$n_N = k_N$$
$$n_{N+1} = 0$$

and

$$k_1 \leq 1$$
$$k_1 + k_2 \leq 2$$
$$\vdots$$
$$\sum_{i=1}^{j} k_i \leq j$$
$$\vdots$$
$$\sum_{i=1}^{N} k_i = N$$

The probability of obtaining the observed results is therefore:

$$P = \sum \frac{n!}{1!1!\cdots1!\,k_1!\cdots k_N!} f_1^{1}\cdots f_K^{1}\, f_1^{k_1}\cdots f_N^{k_N}\, f_{N+1}^{0}$$

$$\left(l \le N \Longrightarrow \sum_{i=1}^{l} k_i \le l,\right) \qquad (3)$$

$$\sum^{N} k_i = N$$

$$= \sum \frac{n!}{k_1!\cdots k_N!}\left[\prod_{i=1}^{K} f(x_i)\,dx\right]\left[\prod_{j=1}^{N}\left\{F(s_{j+1})-F(s_j)\right\}^{k_j}\right]$$

$$\left(l \le N \Longrightarrow \sum_{i=1}^{l} k_i \le l,\right) \qquad (4)$$

$$\sum^{N} k_i = N \qquad \text{where } s_{N+1} = \infty$$

These sums are to be taken over all k such that

$$\sum_{i=1}^{l} k_i \le l \qquad (\text{for all } l \le N)$$

That is, for each term of the sum, all k's must be assigned and they must be assigned such that for all $l \le N$, $\sum_{i=1}^{l} k_i \le l$.

The symbol $\Longrightarrow$ denotes "implies."

The likelihood function is therefore defined:

$$L = \sum \frac{n!}{k_1! \cdots k_N!} \left[\prod_{i=1}^{K} f(x_i) \right] \left[\prod_{j=1}^{N} \left\{ F(s_{j+1}) - F(s_j) \right\}^{k_j} \right]$$

$$\left(l \leq N \Rightarrow \sum_{i=1}^{l} k_i \leq l, \right. \qquad \left(s_{N+1} = \infty \right)$$

$$\left. \sum^{N} k_i = N \right)$$

(5)

This expression may be written logarithmically:

$$L = \sum \frac{n!}{k_1! \cdots k_N!} \exp \left\{ \sum_{i=1}^{K} \ln f(x_i) + \sum_{j=1}^{N} k_j \ln \left[F(s_{j+1}) - F(s_j) \right] \right\}$$

$$\left(l \leq N \Rightarrow \sum_{i=1}^{l} k_i \leq l, \right)$$

(6)

$$\sum^{N} k_i = N$$

Equation (5) may be written more succinctly if one of the products is factored out:

$$L = \prod_{i=1}^{K} f(x_i) \sum_{\substack{\left(l \le N \Rightarrow \sum_{i=1}^{l} k_i \le l,\right. \\ \left. \sum^{N} k_i = N\right)}} \frac{n!}{k_1! \ldots k_N!} \prod_{j=1}^{N} \left\{F(s_{j+1}) - F(s_j)\right\}^{k_j}$$

(7)

$$= S \prod_{i=1}^{K} f(x_i)$$

Consider the case when say $N = 3$. There are five terms for S in this case; the k's for each term are given by the table:

	k_1	k_2	k_3
S_1	1	1	1
S_2	0	2	1
S_3	1	0	2
S_4	0	1	2
S_5	0	0	3

An important special case arises when $s_1 = s_2 = s_3 = \ldots = s_N$. In this case, all terms of S vanish except one; the maximization of the likelihood function is therefore considerably easier. The likelihood function will in general be amenable to maximization only through an approximate numerical method. Such a method may involve taking the first or second partial derivatives of L with respect to the parameters.

REFERENCES:

Shooman, M. L., *Probabilistic Reliability*, McGraw-Hill, pp. 475 - 480

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)

Watervliet Arsenal
Watervliet, N.Y. 12189

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

3. REPORT TITLE

MAXIMUM LIKELIHOOD ESTIMATION OF DISTRIBUTION PARAMETERS FROM GENERALLY CENSORED SAMPLES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

5. AUTHOR(S) (First name, middle initial, last name)

Royce W. Soanes, Jr.

6. REPORT DATE

June 1969

7a. TOTAL NO. OF PAGES

16

7b. NO. OF REFS

1

8a. CONTRACT OR GRANT NO.

-

b. PROJECT NO. -

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

WVT-6923

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

U. S. Army Weapons Command

13. ABSTRACT

A simple derivation of the likelihood function for generally censored samples is given, where "generally censored sample" connotes: a set of failure data containing both failures and survivors with various different tentative survival times.

DD FORM 1 NOV 65 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

Unclassified
Security Classification

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Maximum Likelihood						
Likelihood Function						
Censored Data						
Truncated Tests						

Unclassified

Security Classification

AD ______________ Accession No. ______________
Benet Laboratories, Watervliet Arsenal, Watervliet, N.Y.

MAXIMUM LIKELIHOOD ESTIMATION OF DISTRIBUTION PARAMETERS FROM GENERALLY CENSORED SAMPLES by Royce W. Soanes, Jr.

Report No. WVT-6923, June 1969, 16 pages. Unclassified Report

A simple derivation of the likelihood function for generally censored samples is given, where "generally censored sample" connotes: a set of failure data containing both failures and survivors with various different tentative survival times.

Maximum Likelihood

Likelihood Function

Censored Data

Truncated Tests

Distribution Unlimited

AD ______________ Accession No. ______________
Benet Laboratories, Watervliet Arsenal, Watervliet, N.Y.

MAXIMUM LIKELIHOOD ESTIMATION OF DISTRIBUTION PARAMETERS FROM GENERALLY CENSORED SAMPLES by Royce W. Soanes, Jr.

Report No. WVT-6923, June 1969, 16 pages. Unclassified Report

A simple derivation of the likelihood function for generally censored samples is given, where "generally censored sample" connotes: a set of failure data containing both failures and survivors with various different tentative survival times.

Maximum Likelihood

Likelihood Function

Censored Data

Truncated Teats

Distribution Unlimited

AD ______________ Accession No. ______________
Benet Laboratories, Watervliet Arsenal, Watervliet, N.Y.

MAXIMUM LIKELIHOOD ESTIMATION OF DISTRIBUTION PARAMETERS FROM GENERALLY CENSORED SAMPLES by Royce W. Soanes, Jr.

Report No. WVT-6923, June 1969, 16 pages. Unclassified Report

A simple derivation of the likelihood function for generally censored samples is given, where "generally censored sample" connotes: a set of failure data containing both failures and survivors with various different tentative survival times.

Maximum Likelihood

Likelihood Function

Censored Data

Truncated Tests

Distribution Unlimited

AD ______________ Accession No. ______________
Benet Laboratories, Watervliet Arsenal, Watervliet, N.Y.

MAXIMUM LIKELIHOOD ESTIMATION OF DISTRIBUTION PARAMETERS FROM GENERALLY CENSORED SAMPLES by Royce W. Soanes, Jr.

Report No. WVT-6923, June 1969, 16 pages. Unclassified Report

A simple derivation of the likelihood function for generally censored samples is given, where "generally censored sample" connotes: a set of failure data containing both failures and survivors with various different tentative survival times.

Maximum Likelihood

Likelihood Function

Censored Data

Truncated Teats

Distribution Unlimited

AD ________ Accession No. ________
Benet Laboratories, Watervliet Arsenal, Watervliet, N.Y.

MAXIMUM LIKELIHOOD ESTIMATION OF DISTRIBUTION PARAMETERS FROM GENERALLY CENSORED SAMPLES by Royce W. Soanes, Jr.

Report No. WVT-6923, June 1969, 16 pages. Unclassified Report

A simple derivation of the likelihood function for generally censored samples is given, where "generally censored sample" connotes: a set of failure data containing both failures and survivors with various different tentative survival times.

Maximum Likelihood
Likelihood Function
Censored Data
Truncated Tests

Distribution Unlimited

AD ________ Accession No. ________
Benet Laboratories, Watervliet Arsenal, Watervliet, N.Y.

MAXIMUM LIKELIHOOD ESTIMATION OF DISTRIBUTION PARAMETERS FROM GENERALLY CENSORED SAMPLES by Royce W. Soanes, Jr.

Report No. WVT-6923, June 1969, 16 pages. Unclassified Report

A simple derivation of the likelihood function for generally censored samples is given, where "generally censored sample" connotes: a set of failure data containing both failures and survivors with various different tentative survival times.

Maximum Likelihood
Likelihood Function
Censored Data
Truncated Tests

Distribution Unlimited

AD ________ Accession No. ________
Benet Laboratories, Watervliet Arsenal, Watervliet, N.Y.

MAXIMUM LIKELIHOOD ESTIMATION OF DISTRIBUTION PARAMETERS FROM GENERALLY CENSORED SAMPLES by Royce W. Soanes, Jr.

Report No. WVT-6923, June 1969, 16 pages. Unclassified Report

A simple derivation of the likelihood function for generally censored samples is given, where "generally censored sample" connotes: a set of failure data containing both failures and survivors with various different tentative aurvival times.

Maximum Likelihood
Likelihood Function
Censored Data
Truncated Tests

Distribution Unlimited

AD ________ Acceasion No. ________
Benet Laboratories, Watervliet Arsenal, Watervliet, N.Y.

MAXIMUM LIKELIHOOD ESTIMATION OF DISTRIBUTION PARAMETERS FROM GENERALLY CENSORED SAMPLES by Royce W. Soanes, Jr.

Report No. WVT-6923, June 1969, 16 pages. Unclassified Report

A simple derivation of the likelihood function for generally censored samples is given, where "generally censored sample" connotes: a set of failure data containing both failures and survivors with various different tentative survival times.

Maximum Likelihood
Likelihood Function
Censored Data
Truncated Tests

Distribution Unlimited